

## Claims

What is claimed is:

1. A system for inferring an information goal, comprising:  
a query subsystem adapted to receive at least one of a new query and a new extrinsic data, the query subsystem being operatively coupled to an inference model and a knowledge data store, the query subsystem comprising:  
a natural language processor adapted to parse the new query; and  
an inference engine adapted to infer one or more informational goals based, at least in part, on at least one of the new query, the new extrinsic data and an inference data stored in the inference model.
2. The system of claim 1, where the informational goals include at least one of, a type of information requested in the new query, a topic of the new query, a focal point of the new query, an age of a person presenting a query to the system and one or more levels of detail desired in a response to the new query.
3. The system of claim 2 comprising:  
an input query log adapted to store at least one of, one or more queries and one or more pieces of extrinsic data; and  
a learning system operatively coupled to the input query log, the learning system operable to produce the inference model.
4. The system of claim 3, where the learning system comprises:  
the natural language processor further adapted to produce linguistic data concerning one or more linguistic features;  
a tagging tool adapted to facilitate manipulating the linguistic data;  
one or more taggers adapted to manipulate the linguistic data; and  
the inference model adapted to store information concerning conditional probabilities associated with the likelihood that one or more informational goals exist,

where the conditional probabilities of the informational goals are determined, at least in part, from Bayesian statistical analysis performed on the linguistic data.

5. The system of claim 4, where the linguistic data comprises a parse tree, where the parse tree contains extractable information concerning the nature of and relationships between observable linguistic features.

6. The system of claim 5, where the observable linguistic features in the extractable information comprise word-based features, structural features and hybrid linguistic features.

7. The system of claim 6, where the word-based features indicate the presence of one or more candidate terms that can be employed in predicting an informational goal.

8. The system of claim 4, where the taggers are further adapted to manipulate the linguistic data to conform with one or more schemas associated with reasoning concerning the relevance of a part of a query based on one or more language models.

9. The system of claim 8, where the taggers are further adapted to supervise learning associated with computing probabilities associated with the informational goals.

10. The system of claim 4, where the inference model represents a probabilistic dependency model.

11. The system of claim 4, where the inference model comprises one or more decision trees, the decision trees adapted to store conditional probabilities associated with one or more informational goals, the decision trees being traversable by the linguistic data.

12. The system of claim 3, where the input query log is at least one of a data store and a manual store.

13. The system of claim 3, where the natural language processor is further adapted to parse a new query into one or more parts suitable for retrieving one or more conditional probabilities stored in the inference model.

14. The system of claim 13, where the one or more parts comprise at least one of, logical forms, adjectival phrases, adverbial phrases, noun phrases, verb phrases, prepositional phrases and parse trees.

15. The system of claim 14, the inference engine further adapted to infer one or more informational goals based, at least in part, on at least one of the new query, the new extrinsic data, the one or more parts, and the one or more conditional probabilities stored in the inference model.

16. The system of claim 3, the query subsystem further comprising:  
an answer generator adapted to produce one or more responses to the new query where the one or more responses are based, at least in part, on the inferred informational goals.

17. The system of claim 16, where the one or more responses can vary in at least one of length, precision and detail based, at least in part, on the inferred informational goals.

18. The system of claim 17, where the answer generator is further adapted to produce at least one of an answer responsive to the new query, a rephrased query, a query that can be employed in query by example processing and an error code.

19. The system of claim 16, where the one or more responses are further processed by at least one of a ranking process, a text formatting process, a diagramming process and a text focusing process.

20. The system of claim 1, where the knowledge data store is searchable for information responsive to a new query and where the information retrieved will depend, at least in part, on the inferred informational goals.
21. The system of claim 1 where the query subsystem is compiled into an executable, and where the executable accepts as input one or more query distinctions.
22. The system of claim 2, where informational goals depend, at least in part, on at least one of, the age of a user, the physical location of a user, one or more relationships in which the user is engaged and an application being employed by the user.
23. The system of claim 2, where the level of detail of an answer to a query depends, at least in part, on at least one of, the age of a user, the physical location of a user, one or more relationships in which the user is engaged and an application being employed by the user.
24. The system of claim 1, where the inferred informational goals can be employed to perform at least one of, targeting advertising, link recommendation, age determination, and demographic modeling.
25. The system of claim 2, where one or more inferred informational goals can be employed to adapt at least one of, a guided search process, a guided retrieval process, a post-search filtering process and a new text composition process.
26. The system of claim 1, further comprising an inference clarifier adapted to clarify an inference associated with an informational goal.
27. The system of claim 26 where the inference clarifier further comprises a user interface operable to facilitate clarifying the inference associated with the informational goal.

28. A data packet adapted to be transmitted between two or more computer processes, the data packet comprising:

- linguistic data;
- conditional probabilities; and
- inferred informational goals.

29. A computer readable medium storing computer executable components of a system for inferring an information goal, the system comprising:

- a query component adapted to receive a new query and a new extrinsic data, the query component operatively coupled to an inference model and a knowledge data store, the query component comprising:

- a natural language processing component adapted to parse the new query;
  - and
  - an inference component adapted to infer one or more informational goals based, at least in part, on at least one of, the new query, the new extrinsic data and an inference data stored in the inference model.

30. A method for answering questions comprising:

- inputting a question;
- employing natural language processing to parse the question;
- employing parse data produced by parsing the question to access a decision model, the decision model storing conditional probabilities associated with informational goals;
- inferring one or more informational goals; and
- producing an output related to the question and the one or more inferred informational goals.

31. The method of claim 30, where parsing the question can produce parse data from which at least one of, the existence of relationships between linguistic components in the question, the nature of relationships between linguistic components in the question, parts of speech in the question, logical forms of words in the question, logical forms of phrases

in the question, structural features in the question, the number of distinct points of speech in the question, whether the main noun in the question is singular and whether the question contains a proper noun, can be extracted.

32. The method of claim 30, where accessing the decision model comprises:  
identifying one or more decision trees to access, based at least in part on the parse data;  
traversing the one or more decision trees based, at least in part on the parse data;  
and  
retrieving one or more conditional probabilities from the one or more decision trees.

33. The method of claim 30, where inferring one or more informational goals comprises:  
evaluating the one or more conditional probabilities retrieved from the decision model to determine which, if any, informational goals can be inferred from the one or more conditional probabilities.

34. The method of claim 30, where producing an output related to the question and the one or more inferred informational goals comprises:  
generating at least one of, an answer related to the question and the one or more inferred informational goals, the answer being retrieved from a knowledge base, a rephrased question, the question rephrasing based, at least in part, on the one or more inferred informational goals, one or more sample questions suitable for use in a query by example system and an error code.

35. The method of claim 33, where the inferred informational goals accurately reflect, at least seventy five percent of the time, actual informational goals of a consumer offering a question with less than seven phrases.

36. The method of claim 33, where the inferred informational goals accurately reflect, at least sixty percent of the time, actual informational goals of a consumer offering a question with seven or more phrases.

37. The method of claim 33, where inferring the one or more informational goals further comprises considering at least one of, the age of a user, the physical location of a user, one or more relationships in which the user is engaged and an application being employed by the user.

38. The method of claim 37, where inferring the one or more informational goals further comprises inferring one or more levels of detail for an answer to the question, where inferring the one or more levels of detail for an answer comprises considering at least one of, the age of a user, the physical location of a user, one or more relationships in which the user is engaged and an application being employed by the user.

39. The method of claim 30 further comprising:  
initializing the decision model;  
collecting one or more training sets of questions;  
analyzing the one or more training sets of questions;  
updating the decision model based on the analysis of the one or more training sets of questions; and  
testing the decision model using one or more testing sets of questions.

40. The method of claim 39, where analyzing the one or more training sets of questions comprises:  
parsing questions in the one or more training sets of questions to produce linguistic data concerning the questions;  
employing supervised learning to establish an inference model; and  
employing Bayesian statistical analysis to compute one or more conditional probabilities associated with one or more informational goals based, at least in part, on the linguistic data.

41. The method of claim 40, where updating the decision model comprises:  
updating one or more data structures employed to store conditional probabilities associated with one or more informational goals.
42. The method of claim 41, where updating the one or more data structures comprises:  
adding zero or more nodes to one or more decision trees;  
removing zero or more nodes from one or more decision trees;  
adding zero or more leafs to one or more decision trees;  
removing zero or more leafs from one or more decision trees;  
recomputing zero or more split nodes for one or more decision trees;  
adding zero or more decision trees to the inference model; and  
removing zero or more decision trees from the inference model.
43. The method of claim 41, where testing the decision model comprises:  
producing one or more sets of outputs in response to one or more input testing sets of questions; and  
analyzing the responsiveness of the one or more sets of outputs to the information sought in the one or more input testing sets of questions.
44. The method of claim 30 further comprising adapting one or more of a guided search process, a guided retrieval process, a post-search filtering process and a new text composition process.
45. The method of claim 30, further comprising clarifying an inference associated with an informational goal.
46. A computer readable medium storing computer executable instructions operable to perform a method for answering questions, the method comprising:  
inputting a question;  
employing natural language processing to parse the question;



employing parse data produced by parsing the question to access a decision model, the decision model storing conditional probabilities associated with informational goals;

inferring one or more informational goals; and

producing an output related to the question and the one or more inferred informational goals.

47. A system for answering a question posed to an automated question answerer comprising:

means for initializing a model representing the likelihood that a certain type of answer is desired, the model being stored in one or more data repositories;

means for decomposing a question into parts that facilitate accessing the model;

manual means for adapting the model;

automated means for adapting the model; and

means for retrieving one or more answers to the question based on likelihoods retrieved from accessing the model.

48. A system for learning how to infer information goals from queries, comprising:

a natural language processor adapted to produce a linguistic data concerning one or more linguistic features;

a tagging tool adapted to facilitate manipulating the linguistic data;

one or more taggers adapted to manipulate the linguistic data; and

an inference model adapted to store information concerning conditional probabilities associated with the likelihood that one or more informational goals exist, where the conditional probabilities of the informational goals are determined, at least in part, from Bayesian statistical analysis performed on the linguistic data.

49. The system of claim 48, where the linguistic data comprises a parse tree, where the parse tree contains extractable information concerning the nature of and relationships between observable linguistic features.

50. The system of claim 49, where the observable linguistic features comprise word-based features, structural features and hybrid linguistic features.

51. The system of claim 50, where the word-based features indicate the presence of one or more candidate terms that can be employed in predicting an informational goal.

52. The system of claim 48, where the taggers are further adapted to manipulate the linguistic data to conform with one or more schemas associated with reasoning concerning the relevance of a part of a query based on one or more language models.

53. The system of claim 52, where the taggers are further adapted to supervise learning associated with computing probabilities associated with the informational goals.

54. The system of claim 48, where the inference model comprises one or more decision trees, the decision trees adapted to store conditional probabilities associated with one or more informational goals, the decision trees being traversable by the linguistic data.

55. A computer readable medium storing computer executable components of a system for learning how to infer information goals from queries, the system comprising:

- a natural language processing component adapted to produce a linguistic data concerning one or more linguistic features;
- a tagging component adapted to facilitate manipulating the linguistic data;
- one or more taggers adapted to manipulate the linguistic data; and
- an inference model component adapted to store information concerning conditional probabilities associated with the likelihood that one or more informational goals exist, where the conditional probabilities of the informational goals are determined, at least in part, from Bayesian statistical analysis performed on the linguistic data.

56. A data packet adapted to be transmitted between two or more computer processes,  
the data packet comprising:  
linguistic data;  
conditional probabilities; and  
inferred informational goals.